POINT TARGET DETECTION IN IR IMAGE SEQUENCES USING SPATIO-TEMPORAL HYPOTHESIS TESTING

February 1999

Alexis P. Tzannes and Jonathan M. Mooney
Air Force Research Laboratory
Hanscom AFB, MA, 01731

ABSTRACT

This paper addresses the problem of detecting weak, moving point targets in infrared (IR) image sequences that also contain evolving cloud clutter. The problem is initially attacked in the temporal domain, where there is a clear distinction between targets and cloud clutter. We formulate the temporal detection problem in the context of a hypothesis testing procedure on individual pixel temporal profiles, leading to a theoretically sound and computationally efficient statistical test. The technique assumes we have deterministic and statistical models for the temporal behavior of the background noise, target and clutter, on a single pixel basis. The target temporal profile can be modeled by scaled versions of the point spread function (PSF) of the imager, while the clutter can be well described using a first order Markov model. Based on these models, which are experimentally verified using real data, we develop a generalized likelihood ratio test and perfect measurement performance analysis, and present the resulting decision rule. We demonstrate the effectiveness of the technique by applying the resulting algorithm to real world infrared image sequences containing targets of opportunity. For severe clutter situations which result in false alarms, we suggest an additional spatial hypothesis testing procedure, designed to exploit the difference in the spatial signature of point targets and cloud clutter. As for the temporal case, we propose models for the spatial signatures of targets and cloud clutter and derive the resulting decision rule. Application to real IR image sequences shows that the composite spatio-temporal algorithm results in reduced false alarm rates and increased probability of detection compared to the purely temporal approach.

Form SF298 Citation Data

Report Date *("DD MON YYYY")* 00021999	**Report Type** N/A	**Dates Covered (from... to)** *("DD MON YYYY")*

Title and Subtitle Point Target Detection in IR Image Sequences using Spatio-Temporal Hypotheses Testing	**Contract or Grant Number**
	Program Element Number
Authors	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Air Force Research Laboratory Hanscom AFB, MA, 01731	**Performing Organization Number(s)**
Sponsoring/Monitoring Agency Name(s) and Address(es)	**Monitoring Agency Acronym**
	Monitoring Agency Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited	
Supplementary Notes	
Abstract	
Subject Terms	
Document Classification unclassified	**Classification of SF298** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** unlimited
Number of Pages 19	

1 Introduction

Early work in IR search and track systems utilized algorithms that first attempted to detect the target spatially in each image, and then applied a temporal association or tracking algorithm [5, 8]. Although these algorithms were adequate for early applications in which the targets were bright compared to the background, they performed poorly with dim targets in severe real world clutter [4]. An additional limitation of these algorithms stemmed from the fact that the temporal behavior of the target and clutter was not used to its full extent, since spatial detection and thresholding were performed first.

More recent approaches used multiple frames to incorporate temporal as well as spatial information, they are often referred to as "track before detect" algorithms. The standard approach was to pose the tracking problem as the detection of a known signal in 3-D noise. Assumptions were made on the characteristics of the noise and the optimum linear filter was derived. The initial work performed by Reed *et al.* [15] derived the filter that maximized the output SNR in the general case of noise with known auto-covariance function. For situations in which the clutter in the entire scene did not follow a particular model, partitioning the images into areas with different clutter characteristics was proposed [12]. Alternatively, spatial or spatio-temporal pre-whitening of the images has been performed [6, 3]. A drawback of these track-before-detect techniques is that they are very computationally intensive since the entire 3-D space must be filtered for all possible trajectories for each target velocity. Pre-processing the images, such as pre-whitening add to the computational expense. Suboptimal approaches have been proposed using dynamic programming to reduce the computational complexity [2, 7, 1] but performance was reduced for dim targets in severe clutter.

To summarize, spatial processing of single images followed by association and tracking involves moderate computational complexity but performs poorly for small, weak, moving targets in severe clutter, since the target signal-to-noise ratio (SNR) and signal-to-clutter ratio (SCR) in a single image frame are very low. Full three-dimensional spatio-temporal domain processing provides a higher SNR and SCR domain but such approaches are computationally prohibitive for many practical applications and generally rely on hard clutter assumptions, which have not been shown to be valid for real world clutter. Due to the limited availability of real world image sequences containing real clutter and targets, the majority of the algorithms described above were tested on simulated datasets using embedded targets. As a result, none of these techniques were specifically designed (or have been shown to work successfully) for detecting weak slow targets in scenes that contain severe clutter. To our knowledge, there has not yet been a theoretically solid approach that is relatively computationally efficient and performs well for detecting weak slow targets in the presence of severe real world IR clutter.

The work reported in this paper attempts to fill this void by formulating the detection problem in a hypothesis testing framework, which leads to a computationally efficient detection approach. We believe that the key insight is to process in time **first**, operating on the temporal profile of each pixel. Because the temporal behavior of the target (from a single pixel viewpoint) is distinct from that of the clutter, we can expect that temporal profiles of pixels through which targets pass will be distinct from those through which clutter passes. Therefore the effective temporal SNR and SCR will be higher than the spatial SNR and SCR, resulting in higher probability of detection (P_D) and lower probability of false alarm (P_{FA}) when compared to processing spatially first. In addition, because we are processing each pixel profile independently in the (1-D) temporal domain first, we

achieve a major reduction in computational cost compared to a 3-D hypothesis test.

2 Proposed Approach

The idea of temporal processing (i.e. first processing the profile of each pixel in time) was introduced in earlier work in our laboratory [13, 17]. In these papers the authors presented a heuristic temporal filtering algorithm and demonstrated that temporal processing can be a powerful tool for detecting small moving targets, providing good clutter suppression at relatively low computational cost. In this paper, the insight in [13, 17] is developed in a hypothesis testing framework for target detection and clutter rejection. We develop a two-stage temporal test: the first stage eliminates the vast majority of noise-only pixels and the second stage decides between target and clutter on the remaining pixels. The latter decision rule is somewhat unusual in that under one hypothesis the observed signal is a deterministic signal with unknown parameters in noise, while under the other it is a random signal in noise. For difficult clutter situations which may cause false alarms in the temporal test, we develop a subsequent spatial hypothesis test. The spatial test is performed only on those pixels that were above the threshold after both temporal stages.

3 Temporal Processing

In this section we present a temporal approach to the detection of point targets in image sequences using a hypothesis testing formulation. In Sec. 3.1 we develop statistical models for the temporal profiles of pixels that see clear sky, targets and cloud clutter. These models are then utilized in Sec. 3.2 where we develop and analyze the corresponding 3-ary hypothesis testing procedure. A computationally efficient suboptimal approach is described in Sec. 3.3.

3.1 Pixel Temporal Profile Modeling

Temporal processing exploits the difference between the temporal profiles of pixels through which a target passes, compared to those affected only by clear sky and those affected only by cloud clutter. In this section we introduce deterministic and statistical models for the temporal profiles of targets, clutter and clear sky. These models have been developed through the study of real IR cameras and a large database of real IR image sequences. The performance of these IR cameras is similar to those previously characterized [14].

Pixels that see clear sky or other features which are constant in time will have time profiles that generally behave like a constant mean value plus white noise. Stationary or very large slow moving clutter will also appear as a slowly varying mean plus the same background noise process. A pixel that is affected by a small moving target will have a pulse-like shape similar to a scaled version of the PSF of the camera. The width and height of the pulse are related to the target velocity and intensity, respectively. Pixels that are affected by cloud edges or other difficult clutter features will have temporal profiles that behave less regularly. As we will show below, these pixels can be modeled using a first order Markov model. In the following three sections we elaborate on these characteristics of pixel temporal profiles and develop a model for each type. These models are then used in the derivations of Sec. 3.2 where we pose the temporal target detection problem in the context of a hypothesis testing procedure.

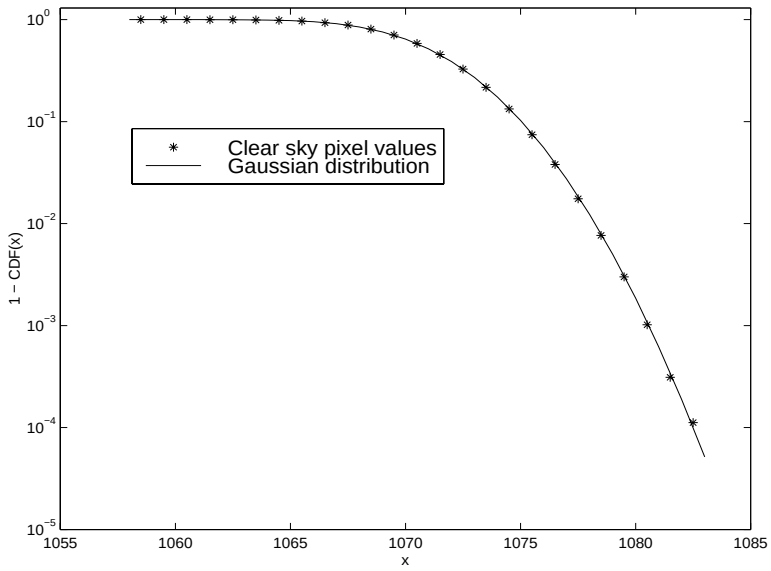

Figure 1: Comparison of 1 minus the CDF of clear sky pixel values to a Gaussian distribution function with mean and variance equal to that of the empirical data.

3.1.1 Clear Sky Pixel Model

The temporal profiles of pixels that see clear sky or other features that are constant in time can be modeled by the sum of a constant and a temporally white Gaussian background noise term. The Gaussian noise term has zero mean and a standard deviation that is constant for a given camera, usually between 3 and 5 analog to digital converter units (ADU). Denoting a clear sky pixel time profile as $p_{cs}(k)$, where k is a sampled time variable, we can model these pixels as

$$p_{cs}(k) = C + n(k), \qquad n(k) \sim \mathcal{N}(0, \sigma_n) \quad \text{with} \quad E\{n(k)n(k+m)\} = \sigma_n^2 \delta(m), \tag{1}$$

where C is a constant, $n(k)$ is the background noise term and $\delta(k)$ is the Kronecker delta sequence.

The validity of this model is illustrated in Fig. 1. This figure displays the empirical CDF of the pixel values of a small area in an image sequence that is seeing clear sky. The values of 100 pixels over a time interval of 3 seconds (90 frames) were used to create the histogram. Also shown on the plot for comparison is a Gaussian CDF with the same mean and standard deviation as the empirical data. Note that we actually plot 1 minus the CDF so that the resolution of the tails is better on a semi-log plot. It is apparent from this figure that the Gaussian assumption is reasonable, as the empirical CDF of the pixel values closely matches the Gaussian CDF. To illustrate that this Gaussian noise can be modeled as white, in Fig. 2 we have plotted the Power Spectral Density (PSD) of a single pixel in time, consisting of 1500 samples (50 seconds). Notice that the response is flat over all frequencies, except at DC, indicating that the process is white.

3.1.2 Target Pixel Model

The time signature of a pixel affected by a small moving target will have a pulse-like shape caused by the target moving across the pixel. The width of the pulse is inversely proportional to the speed

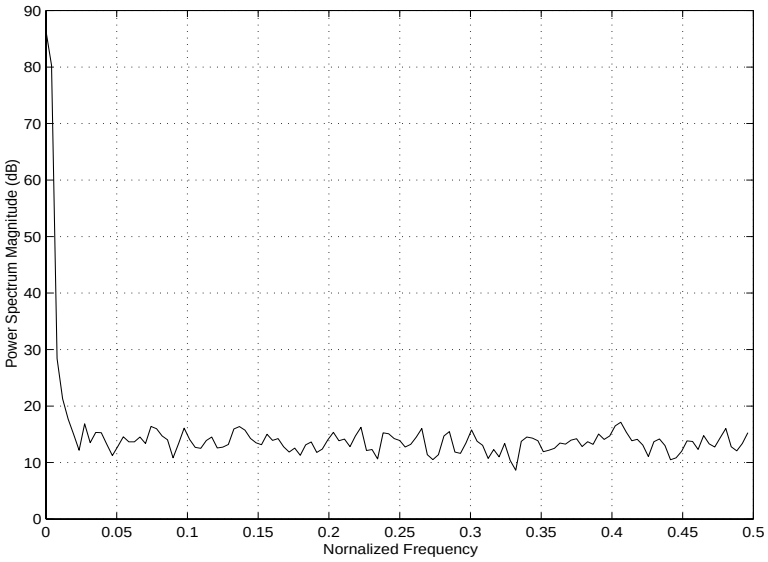

Figure 2: PSD of the temporal profile of a single pixel in time, consisting of 1500 samples. Notice that the PSD is relatively constant over all frequencies, (except for the DC term at $f = 0$), indicating that the sequence is white.

of the target, whereas the intensity of the pulse is proportional to the its strength. If the target is modeled as a moving point source, the target profiles can then be modeled as dilated or contracted versions of one dimensional profiles of the PSF of the imager. Thus if the imager PSF is known, we have a known deterministic model for the target signal with unknown parameters which depend on target velocity, intensity, and time of arrival, and on the background level. Denoting this PSF profile as $f(k; \mathbf{p})$, where k is the sampled time variable and $\mathbf{p}$ is the unknown parameter vector, the target model can be written as

$$p_{tar}(k) = f(k; \mathbf{p}) + n(k), \tag{2}$$

where $n(k)$ is the background noise term introduced in Eq. 1.

In previous work in our laboratory [20], we developed a technique to measure the PSF of the staring IR cameras used to acquire the image sequences used in this work. These measurements showed that the derivative of a Fermi function is a suitable choice for the pulse-like function $f(k; \mathbf{p})$ in Eq. 2. The derivative of the Fermi function is given by

$$f_{df}(k; a, b, c, d) = \frac{a \exp\left[(k-b)/c\right]}{\{\exp\left[(k-b)/c\right] + 1\}^2} + d. \tag{3}$$

The parameter a is proportional to the intensity of the target, while b determines the time of arrival, i.e. the time instant at which the target is centered on the pixel. The parameter c is a scale parameter that determines the width of the function and d is the background level.

To illustrate, in Fig. 3 we show five measured target temporal profiles extracted from real IR image sequences in the left column alongside profiles simulated using our model in the right column. The parameters in $\mathbf{p}$ were varied to simulate targets of different velocities, intensities, time of arrival and background level.

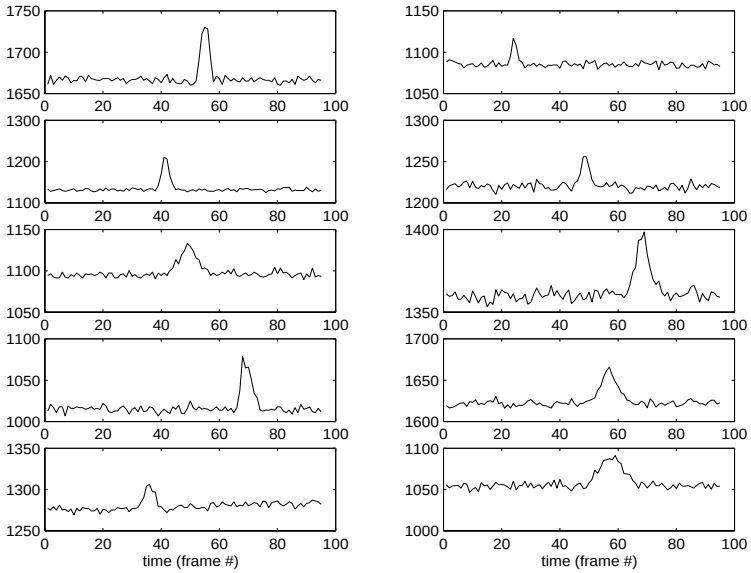

Figure 3: Plots of temporal pixel profiles of real targets extracted from IR sequences shown in left column. Simulated targets, created using the model in Eq. 2 with varied values of the unknown parameters, are shown in the right column.

3.1.3 Cloud Clutter Pixel Model

Pixels in the scene that see cloud clutter have temporal profiles that behave less regularly. The bright edges of clouds cause these pixels to have peaks that are broader than those seen in target profiles. A set of ten cloud clutter pixel profiles extracted from one of our IR image sequences is shown in Fig. 4. By analyzing the correlation structure of clutter profiles extracted from our sequences we concluded (for details, see [18]) that a simple first order AR or random walk model is suitable for these profiles. For a given clutter pixel, the value of the pixel at time k can be expressed as the sum of the pixel value at time $k-1$ and a Gaussian noise or error term. Denoting the value of a clutter pixel as $p_{cl}(k)$, we have

$$p_{cl}(k) = p_{cl}(k-1) + w(k), \qquad \text{where} \;\; w(k) \sim \mathcal{N}(0, \sigma_c). \tag{4}$$

The Gaussian background noise $n(k)$ has been incorporated in the $w(k)$ term. The magnitude of the parameter σ_c, which is the standard deviation of the driving noise of the model, describes the severity of the clutter in a scene.

We verified that the clutter time profiles follow this model by investigating the statistics of the first order temporal differences $p_{cl}(k) - p_{cl}(k-1)$. In Fig. 5 we show the empirical CDF of the first order temporal differences of clutter pixels extracted from a sample image sequence. To create the CDF we used the temporal differences of 1500 pixels over 95 frames. Note that again, as in Fig. 1, we actually plot 1 minus the CDF so that the resolution of the tail is better on a semi-log plot. A Gaussian CDF with the same mean and standard deviation as the empirical data is drawn on the plot for comparison.

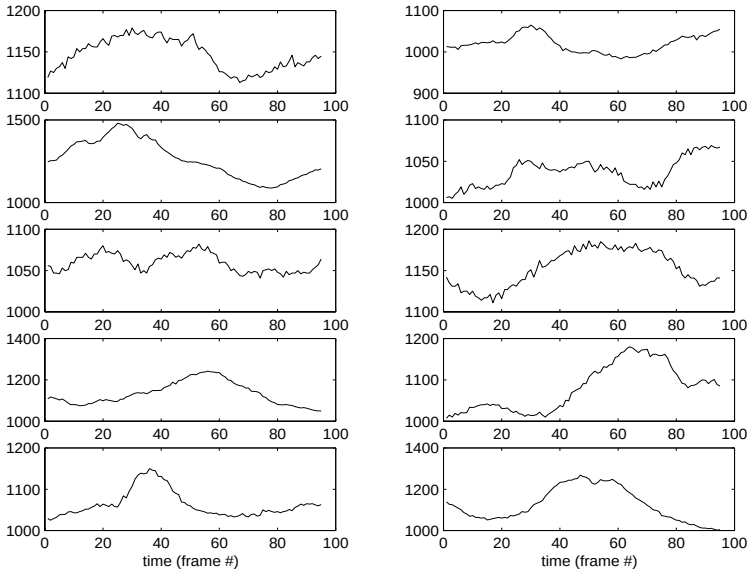

Figure 4: Plots of temporal profiles of cloud clutter pixels.

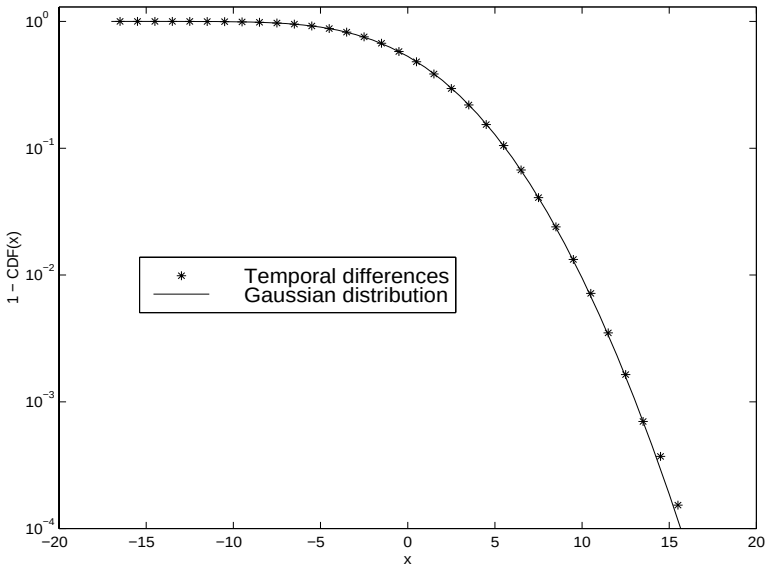

Figure 5: Comparison of 1 minus the CDF of first order temporal and variance equal to that of the empirical data.

3.1.4 Definitions of SNR and SCR

Using parameters from the models introduced in the previous sections, we now define two metrics that relate target strength to the background noise and cloud clutter in the scene. The definitions of the target SNR and temporal SCR presented in this section will be used in the remainder of the paper. For an ideal target temporal profile, we define *target intensity*, denoted I_t, as the highest deviation of any single signal sample from the background level (the parameter d in Eq. 3) over the entire sequence of length N. Target SNR is then defined as

$$SNR = \frac{I_t}{\sigma_n}, \tag{5}$$

where σ_n is the background noise standard deviation.

As mentioned above, the standard deviation σ_c of the driving noise of the Markov model is a measure of the severity of the cloud clutter in a sequence. The value of σ_c is relatively constant over all clutter pixels in a given image sequence. Therefore, we define the target temporal SCR, denoted SCR as

$$SCR = \frac{I_t}{\sigma_c}. \tag{6}$$

3.2 3-ary Hypothesis Testing Formulation

With the models presented in the previous section, the temporal detection problem leads to a 3-ary hypothesis testing scenario. An observed temporal profile, which we denote $r(k)$, consists of constant plus noise under H_0, cloud clutter under H_1 and target plus noise under H_2.

$$\begin{aligned} H_0 : r(k) &= p_{cs}(k) = C + n(k) \\ H_1 : r(k) &= p_{cl}(k) = r(k-1) + w(k) \\ H_2 : r(k) &= p_{tar}(k) = f(k;\mathbf{p}) + n(k) \end{aligned}$$

Denoting the received signal vector of length N as $R = [r(1), r(2), ..., r(N)]^T$, the likelihood function of R under the assumption that H_0 is true, denoted $p_0(R)$, is

$$p_0(R) = \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left[\frac{-[r(k) - C]^2}{2\sigma_n^2}\right]. \tag{7}$$

The likelihood function of R under the assumption that H_1 is true, denoted $p_1(R)$, can be expressed as [10]

$$p_1(R) = \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi\sigma_c^2}} \exp\left[\frac{-[r(k) - r(k-1)]^2}{2\sigma_c^2}\right], \tag{8}$$

since the temporal differences $r(k) - r(k-1)$ follow a Gaussian distribution with mean zero and standard deviation σ_c. The likelihood function of R under the assumption that H_2 is true, denoted $p_2(R)$, is given by

$$p_2(R) = \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left[\frac{-[r(k) - f(k;\mathbf{p})]^2}{2\sigma_n^2}\right], \tag{9}$$

since the received signal samples $r(k)$ will be IID Gaussian random variables with mean $f(k;\mathbf{p})$ and standard deviation σ_n.

In a 3-ary hypothesis testing scenario one constructs three different likelihood ratios using the three pairs of likelihood functions, and tests each against a threshold to decide which hypothesis is more likely. We are not interested in discriminating between noise and clutter pixels, since our goal is to detect target pixels. Therefore, the overall structure becomes simpler if we first use the following log-likelihood ratio, denoted $\lambda_1(R)$, to separate target or clutter from noise or clutter:

$$\lambda_1(R) = \ln\left[\frac{p_2(R)}{p_0(R)}\right] \begin{array}{c} H_2 \text{ or } H_1 \\ > \\ < \\ H_0 \text{ or } H_1 \end{array} T_1. \tag{10}$$

If this ratio is above the threshold T_1 we decide that the signal is either target or clutter, and proceed to a second test with the log-likelihood ratio

$$\lambda_2(R) = \ln\left[\frac{p_2(R)}{p_1(R)}\right] \begin{array}{c} H_2 \\ > \\ < \\ H_1 \end{array} T_2, \tag{11}$$

to decide whether the profile R was a target pixel or a cloud clutter pixel. If the first ratio $\lambda_1(R)$ is below the threshold T_1 we decide that the profile R is either clear sky or clutter, and no further processing is required since we are not interested in distinguishing between these two hypotheses.

Using the expressions for the likelihood functions $p_0(R)$ and $p_2(R)$ from Eq. 7 and 9, the log-likelihood ratio $\lambda_1(R)$ can be written as

$$\lambda_1(R) = \sum_{k=1}^{N}\left[\frac{[r(k)-C]^2}{2\sigma_n^2} - \frac{[r(k)-f(k;\mathbf{p})]^2}{2\sigma_n^2}\right] \begin{array}{c} H_2 \text{ or } H_1, \text{ proceed to } \lambda_2(R) \\ > \\ < \\ H_0 \text{ or } H_1, \text{ done} \end{array} T_1 \tag{12}$$

which can be simplified to be

$$\lambda_1(R) = \frac{1}{2\sigma_n^2}\left[\sum_{k=1}^{N} 2r(k)[f(k;\mathbf{p})-C] + C^2 - f(k;\mathbf{p})^2\right]. \tag{13}$$

Similarly, using the expressions for $p_1(R)$ and $p_2(R)$ from Eq. 8 and 9 the second log-likelihood ratio $\lambda_2(R)$ can be written as:

$$\lambda_2(R) = \ln\left[\frac{\sigma_c^N}{\sigma_n^N}\right] + \sum_{k=1}^{N}\left[\frac{[r(k)-r(k-1)]^2}{2\sigma_c^2} - \frac{[r(k)-f(k;\mathbf{p})]^2}{2\sigma_n^2}\right]. \tag{14}$$

We can absorb the $\ln\left[\frac{\sigma_c^N}{\sigma_n^N}\right]$ term in Eq. 14 into the threshold T_2 because it does not depend on the specific temporal profile R (since the standard deviations of both the background noise and the clutter model are assumed constant for a given image sequence), leading to the modified test:

$$\sum_{k=1}^{N}\left[\frac{[r(k)-r(k-1)]^2}{2\sigma_c^2} - \frac{[r(k)-f(k;\mathbf{p})]^2}{2\sigma_n^2}\right] \begin{array}{c} H_2 \\ > \\ < \\ H_1 \end{array} T_2. \tag{15}$$

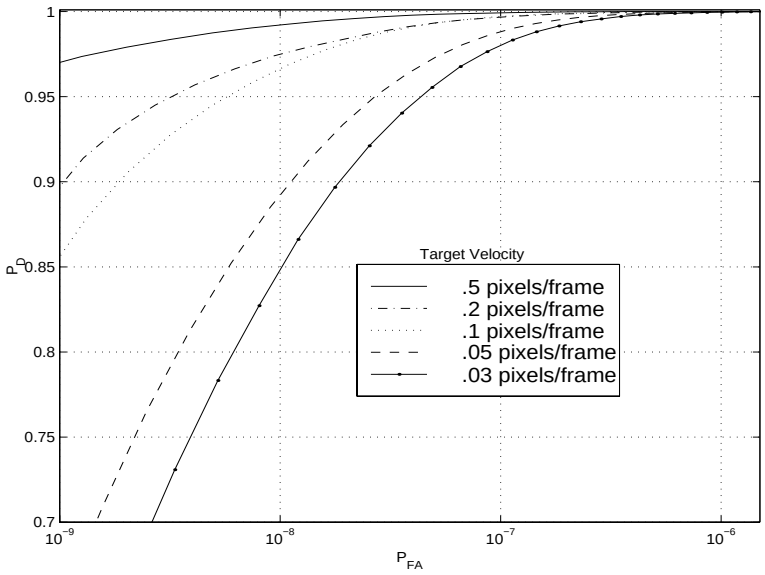

Figure 6: Final ROC curves for targets of varying velocities, at fixed SNR = 4 and SCR = 2.67.

We note that the detection statistic of Eq. 15 consists of two summations: The first, which tests how well the profile R matches the clutter model, is the energy of the first order temporal differences of R normalized by the variance σ_c^2. This term will be small for temporal profiles that match the clutter model and large for profiles that do not. The second, which tests how well R matches the target model, is simply the mean square error between R and the target profile $f(k;\mathbf{p})$, normalized by the background noise variance σ_n^2. This term will be small for temporal profiles that match the target profile and large for profiles that do not. The algebraic sum of these two terms will result in $\lambda_2(R)$ being large for target profiles and small (or negative) for clutter profiles, as desired for the decision rule of Eq. 15.

3.2.1 Performance Analysis of 3-ary test

In this section we present a *perfect measurement* ROC curves [21] for the performance of the 3-ary temporal test by assuming knowledge or perfect estimates of the unknown parameters in $\mathbf{p}$, C, σ_c and σ_n. In practice, the unknown parameters would be estimated from the observed signal, causing a degradation in the performance of the test. By deriving the PDF of $\lambda_1(R)$ and $\lambda_2(R)$ under each of the three hypotheses we can create performance analysis plots curves for each of the two tests, which can be used to create ROC curves for the entire test (for details of the derivations and analysis of the performance of $\lambda_1(R)$ and $\lambda_2(R)$, see [18]).

Examples of final ROC curves for the entire 3-ary test for targets of varying velocities are shown in Fig. 6. In creating these curves we assumed values for the unknown parameters that are typical for our image sequences. The target SNR and SCR were set to 4 and 2.67, respectively. As is evident from these ROC curves, slower targets are more difficult to detect than faster targets. This is true because the temporal profiles of slower targets have smaller temporal differences and essentially "look more like" clutter. Also note that even for slow targets the overall $\mathrm{P_{FA}}$ is less than 1×10^{-6} for $\mathrm{P_D}$ values very near 1, corresponding to less than one false alarm per image sequence.

3.3 Suboptimal Alternative to $\lambda_1(R)$

In this section we describe a suboptimal alternative to the first log-likelihood ratio $\lambda_1(R)$. Because the computation of $\lambda_1(R)$ involves the estimation of time of arrival and scale parameters, it is not computationally attractive to compute $\lambda_1(R)$ for every pixel in the sequence. By analyzing the performance of $\lambda_1(R)$, we noticed that it basically eliminates clear sky pixels and allows most of the target pixels and approximately 40% of the clutter pixels to pass to the next stage of the 3-ary test. Therefore we sought a computationally simple alternative to $\lambda_1(R)$ that will serve as pre-processing step to the second log-likelihood ratio $\lambda_2(R)$. Ideally, after this pre-processing step, the remaining pixels can be reliably considered to be either target or cloud clutter pixels. In earlier work [19], we showed that this can be accomplished by applying a temporal bandpass filter to each pixel and thresholding the maximum filter response over the N time samples. After this pre-processing step, the remaining pixels will include pixels that are actual targets as well as cloud clutter pixels that "look like" target pixels. The second log-likelihood ratio $\lambda_2(R)$ can then be applied to differentiate actual targets from cloud clutter pixels.

By analyzing the performance of the bandpass filtering pre-processing operation and comparing it to $\lambda_1(R)$ we concluded that the performance of the bandpass filtering operation was comparable to that of $\lambda_1(R)$. Both operations passed most target pixels and eliminated most the clear sky pixels. Furthermore, the bandpass filtering operation passed fewer clutter pixels to the second stage, approximately 20%. In exchange, the computational complexity of the first stage of the temporal test is reduced by approximately one order of magnitude.

4 Spatial Processing

In the previous section, the problem of detecting targets in a three dimensional space was approached as a one dimensional problem by modeling and testing the temporal profiles of individual pixels. Although assumptions about the spatial characteristics of targets and clutter are inherent in the temporal pixel profile models, the temporal test itself does not directly utilize the spatial signature of the targets and cloud clutter. For those pixels that were above the threshold in the second stage of the temporal test, additional spatial processing may be useful to improve the overall P_D and lower the overall P_{FA}. Towards this goal, in this section we develop a spatial hypothesis testing procedure designed to differentiate between small targets and cloud clutter. The spatial processing algorithm is applied only to those pixels that were above the threshold after the second stage of the temporal test. The spatial test is applied to an $M \times N$ spatial profile centered at the pixel of interest, extracted from the image in the sequence corresponding to the estimated time of arrival of the potential target in the temporal test. A key advantage to performing spatial processing at this stage of the detection procedure is that we have identified a small number of pixels that will be processed spatially, eliminating the need to process entire images with a spatial algorithm. In addition, not all images in the sequence must be processed, since the temporal approach has estimated a time of arrival parameter for potential target pixels, indicating the corresponding image in the sequence where we suspect the target might be present. To derive the spatial hypothesis test we proceed in a manner similar to the temporal test. We first propose spatial models for the targets and cloud clutter and then use these models to derive the corresponding likelihood ratio test.

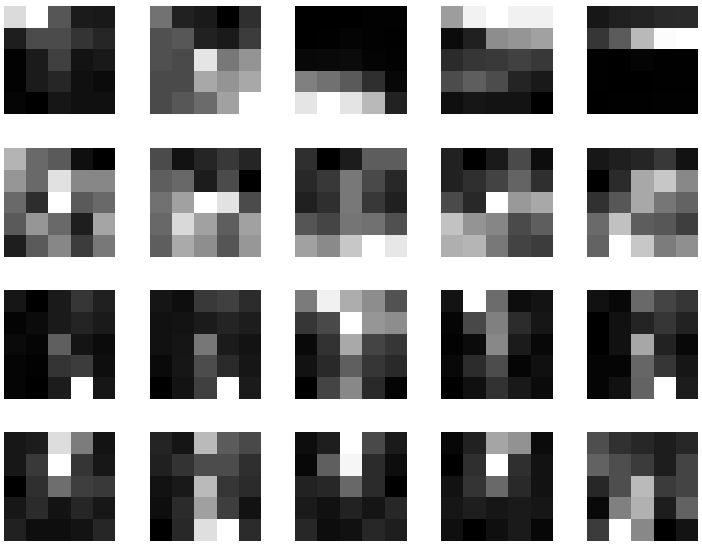

Figure 7: Examples of 5×5 spatial cloud clutter profiles, extracted from a single image sequence.

4.1 Clutter and Target Spatial Profile Modeling

As mentioned above, the spatial hypothesis testing procedure described in this section is applied only to those pixels that were above the threshold in the second stage of the temporal test. Thus, the pixels that reach this spatial test can reliably be considered to be either cloud clutter or target pixels.

4.1.1 Cloud Clutter Spatial Model

The clutter pixels that reach the second stage of the temporal detection algorithm are generally caused by the edges of moving and evolving cloud clutter in the scene. Examples of such profiles are shown in Fig. 7. Notice that these clutter profiles are often quite thin spatially, leading to temporal profiles that resemble those of target pixels. These type of clouds may be expected to cause several potential false alarms after the temporal test.

Because the clouds in our sequences are moving slowly across the scene (less than one pixel/frame), a 1-D spatial profile of a cloud edge is essentially a subsampled version of the temporal profile of the pixel that is seeing the same cloud edge. Therefore we expect that a Markov model should fit the spatial signature of cloud clutter over small spatial regions. More specifically, we investigated the use of a two-dimensional AR model of the form

$$s(m,n) = \sum_{l=-L/2}^{L/2} \sum_{k=-K/2}^{K/2} a_{l,k} s(m-l, n-k) + \epsilon(m,n), \tag{16}$$

where $s(m,n)$ is the cloud clutter spatial profile, $\epsilon(m,n)$ is the two-dimensional driving noise of the model and the $a_{l,k}$ are the AR model coefficients, with $a_{0,0} \stackrel{\triangle}{=} 0$. We assume that $\epsilon(m,n)$ is white Gaussian with variance σ_ϵ^2. The coefficients $a_{l,k}$ can be evaluated by solving the set of linear equations known as the normal equations [9].

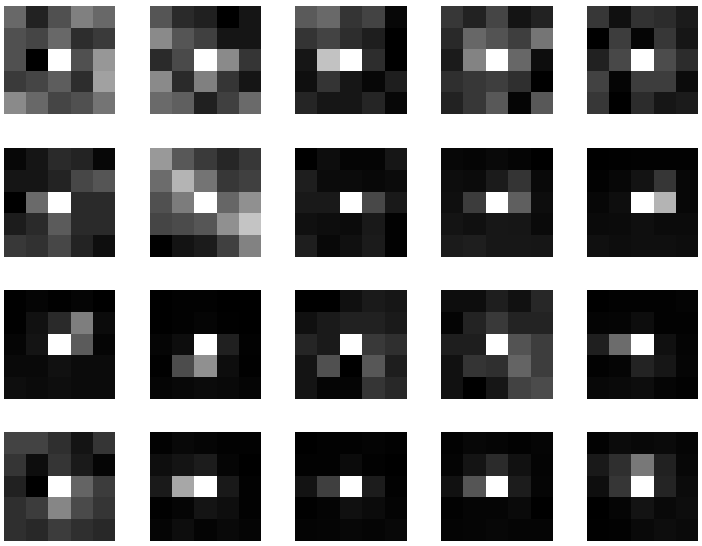

Figure 8: Examples of spatial target profiles extracted from several image sequences.

4.1.2 Target Spatial Model

Assuming that the target is very small, the spatial profile of the target will be a scaled version of the PSF of the imager. Therefore, as in the temporal case, if the PSF of the imager is known, we have a known deterministic model for the target spatial signature with unknown parameters which depend on target size, intensity, exact location on the focal plane and background level. Several examples of target spatial profiles extracted from our IR image sequences are shown in Fig. 8. As with the spatial clutter profiles shown earlier, these target profiles were created by extracting a 5×5 spatial sub-image from the image corresponding to the estimated time of arrival of the target in the calculation of $\lambda_2(R)$ for the temporal profile of the center pixel.

Denoting the PSF model $f(m, n; \mathbf{u})$, where m and n are the sampled space variables and $\mathbf{u}$ is the unknown parameter vector, the target model can be written as

$$s_{tar}(m, n) = f(m, n; \mathbf{u}) + n_s(m, n), \tag{17}$$

where $n_s(m, n)$ is a white Gaussian background noise term with variance σ_n^2.

As mentioned earlier, previous work [20] showed that the 1-D horizontal and vertical profiles of the two-dimensional PSF of our cameras can be modeled using the derivative of the Fermi function. Assuming a circularly symmetric PSF we can interpolate and arrive at a two-dimensional PSF model, which can be used for $f(m, n; \mathbf{u})$ in the target model of Eq. 17.

Using these models for the spatial profiles of targets and clutter we can derive a likelihood ratio test in the same manner as we did for the temporal case. In the interest of space we will not include details of the spatial test in this paper and refer the reader to [18].

5 Results

In this section we demonstrate the effectiveness of the proposed approach by testing it on real IR image sequences containing targets of opportunity and evolving cloud clutter. The sequences were acquired using PtSi IR cameras with 320×244 pixel focal plane arrays. The image data from the camera was captured to 12-bit precision at 30 frames per second using an Ampex digital cassette recorder. Selected sequences consisting of 95 consecutive frames were used for algorithm evaluation. The targets in these sequences are airplanes flying across the scene at long range. We will present detailed results of applying the algorithm to a sample image sequence along with a summary of results for the total of 15 selected sequences.

The proposed approach is a three stage process. First, the temporal profile of each pixel in the sequence is processed using the bandpass filtering alternative to $\lambda_1(R)$. After thresholding the maximum output over the 95 frames, the remaining pixels can reliably be considered to be either target or cloud clutter pixels. These pixels are processed using $\lambda_2(R)$, the second stage of the temporal test described in Sec. 3.2. This test discriminates between target and clutter pixels. The $\lambda_2(R)$ values are thresholded, and for those values that are above the threshold, spatial profiles are extracted from the frame corresponding to the estimated time of arrival of the target in the calculation of $\lambda_2(R)$. The spatial test is applied to these spatial profiles. The pixels above the threshold after the spatial test are declared targets.

5.1 Sample Sequence

The sample image sequence we selected is a daytime scene that included a target and a large amount of drifting and evolving clouds. A sample image from the sequence is shown in Fig. 9. The outlined area indicates the location of the target trajectory. The difference between the pixel values in the first and last frames is shown in Fig. 10. The temporal standard deviation of each pixel over the 95 frames are shown in Fig.11. As is evident from Fig. 10 and 11, the cloud clutter in this scene is quite severe, with large areas of clouds that are changing in time.

After the pre-processing step there were 4327 pixels (5.5%) that were above the threshold. A binary image indicating the location of the pixels that passed the pre-processing step is shown in Fig. 12. Notice the outline of the cloud clutter and the streak in the middle of the image caused by the target pixels. After applying $\lambda_2(R)$ to these pixels and thresholding, only 12 pixels were above the threshold. A binary image showing the location of the pixels that were above the threshold is shown in Fig. 13. Notice that the target detections appear as a streak in the middle of the image, and there are a total of 3 potential false alarms pixels, caused by the cloud clutter. After applying the spatial test, however, these potential false alarms are eliminated, leaving only the correct target detections, as shown in Fig. 14.

5.2 Summary of Results

The algorithm was applied to 15 image sequences, containing a total of 21 targets. After the temporal test, 19 of the 21 targets were detected with a total of 25 areas of potential false alarms caused by the cloud clutter. Applying the spatial test eliminated the potential false alarms in all but 2 of the sequences. So, overall, after the spatial test, we detected 19 out of 21 targets with just 2 areas of false alarms, over the 15 image sequences. Both of these false alarms were caused by

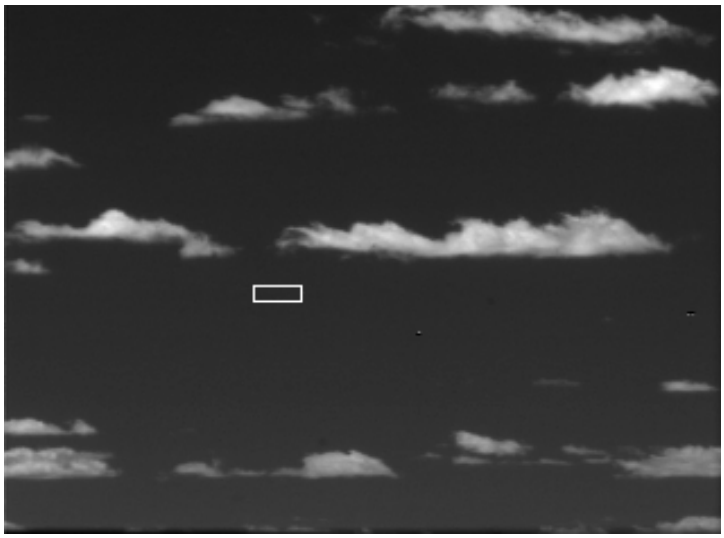

Figure 9: Single image from sample sequence. The outlined area designates the location of the target trajectory.

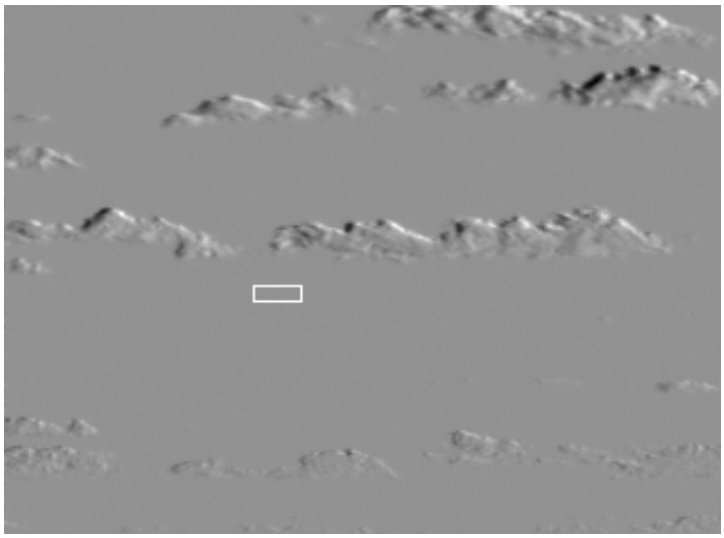

Figure 10: Difference image between the first and last frames in the sequence.

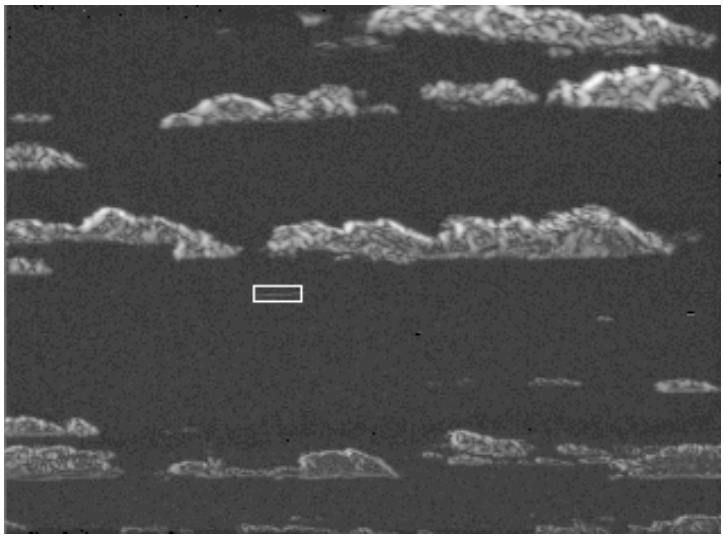

Figure 11: Image of temporal standard deviation of each pixel, over the 95 frame sequence.

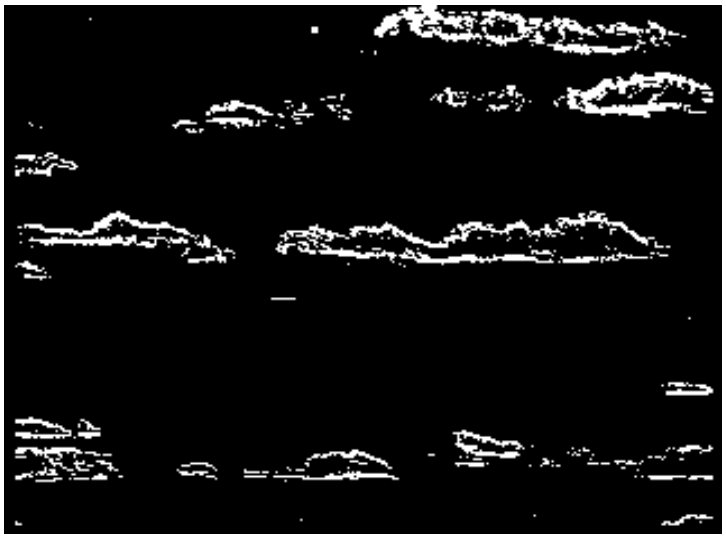

Figure 12: Binary images indicating in white the location of the pixels that passed the pre-processing step.

Figure 13: Binary image indicating in white the location of the pixels that were above the threshold after applying $\lambda_2(r)$.

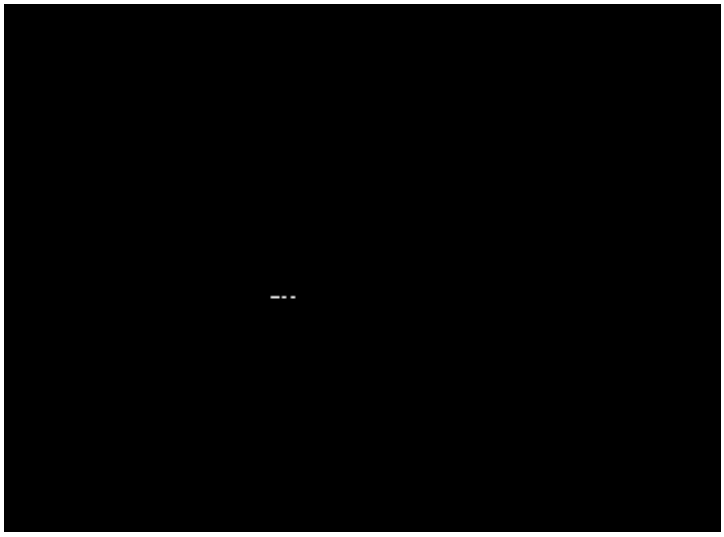

Figure 14: Binary image indicating in white the location of the pixels that were above the threshold after applying the spatial test.

cloud clutter that resembled a target both temporally and spatially. Some possible improvements to the algorithm to eliminate false alarms in these situations are discussed in the following section.

6 Conclusions and Future Work

This paper addresses the problem of detecting small, moving, low amplitude targets in IR image sequences that also contain moving and evolving cloud clutter. We develop a theoretically sound approach to this problem using a hypothesis testing procedure that exploits the difference between the temporal and spatial behavior of targets and clutter. The first stage of the approach - the temporal processing stage - uses experimentally verified temporal models for the clear sky, targets and clutter to develop a temporal hypothesis test that identifies potential target pixels in the image. These pixels are then processed using the second stage of the approach - the spatial processing stage - which uses spatial target and clutter models to develop a similar hypothesis testing procedure designed to further discriminate between targets and clutter. The approach is shown to perform reliably for real world IR image sequences containing drifting and evolving cloud clutter and airplanes flying at long range. In very rare situations, spatially thin fast-moving clouds caused false alarms because of the similarity of such clutter to small targets (both in time and in space). Some possible improvements could be made to the spatial stage of the algorithm to address these situations. Morphological approaches [11, 16] using erosions and dilations could be used to identify these situations and eliminate potential false alarms. Another possibility would be to use a large spatial window and perform an edge detection technique to identify areas of thin wispy clouds. Since targets are not expected to have a spatial extent larger than a few pixels in any direction, we could eliminate areas where edges had a spatial extent larger than a few pixels.

References

[1] J. Arnold and J. Pasternack. Detection and tracking of low-observable targets through dynamic programming. *Proc. SPIE*, 1305:206–217, 1990.

[2] Y. Barniv. Dynamic programming solution for detecting dim moving targets. *IEEE Transactions on Aerospace and Electronic Systems*, AES-21(6):144–156, 1985.

[3] S. D. Blostein and T. S. Huang. Detecting small, moving objects in image sequences using sequential hypothesis testing. *IEEE Transactions on Signal Processing*, SP-19(7):1611–1629, 1991.

[4] D. S. K. Chan. A unified framework for IR target detection and tracking. *Proc. SPIE*, 1698:66–76, 1992.

[5] D. S. K. Chan, D. A. Langan, and D. A. Staver. Spatial processing techniques for the detection of small targets in IR clutter. *Proc. SPIE*, 1305:53–62, 1990.

[6] J. Y. Chen and I. S. Reed. A detection algorithm for optical targets in clutter. *IEEE Transactions on Aerospace and Electronic Systems*, AES-23(1):46–59, 1987.

[7] M. F. Fernandez, A. Aridgides, and D. Bray. Detecting and tracking low-observable targets using IR. *Proc. SPIE*, 1305:193–206, 1990.

[8] C. F. Ferrara, R. W. Fries, W. B. Rushnow, and H. H. Mansur. Adaptive signal processing for the detection of point targets in non-stationary clutter. *Proc. of the IRIS Specialty Group on Targets, Backgrounds and Discrimination*, II:187–198, 1988.

[9] A. K. Jain. *Fundamentals of Digital Image Processing.* Prentice Hall, Englewood Cliffs, NJ, 1989.

[10] D. Kazakos and P. Papantoni-Kazakos. *Detection and Estimation.* Computer Science Press, New York, 1990.

[11] P. Maragos and R. W. Schafer. Morphological systems for miltidimensional signal processing. *Proceedings of the IEEE*, 78(4):690–710, April 1990.

[12] K. A. Melendez and J. W. Modestino. Spatiotemporal multiscan adaptive matched filtering. *Proc. SPIE*, 2561:51–65, 1995.

[13] J. M. Mooney, J. Silverman, and C. E. Caefer. Point target detection in consecutive frame infrared imagery with evolving cloud clutter. *Optical Engineering*, 34(9):2772–2784, 1995.

[14] J. E. Murguia, J. M. Mooney, and W. S. Ewing. Evaluation of a PtSi infrared camera. *Optical Engineering*, 29(7):786–794, (1990).

[15] I. S. Reed, R. M. Gagliardi, and H. M. Shao. Application of three-dimensional filtering to moving target detection. *IEEE Transactions on Aerospace and Electronic Systems*, AES-19(6):898–905, 1982.

[16] J. Rivest and R. Fortin. Detection of dim targets in digital infrared imagery by morphological image processing. *Optical Engineering*, 35(7):1886–1893, 1996.

[17] J. Silverman, J. M. Mooney, and C. E. Caefer. Temporal filters for detecting weak slow point targets in evolving cloud clutter. *IR Physics and Technology*, 37:695–710, 1996.

[18] A. P. Tzannes. *Detection of small targets in infrared image sequences containing evolving cloud clutter.* PhD thesis, Northeastern University, 1998.

[19] A. P. Tzannes and D. H. Brooks. Temporal filters for point target detection in IR imagery. *Proc. SPIE*, 3061:508–520, 1997.

[20] A. P. Tzannes and J. M. Mooney. Measurement of the modulation transfer function of infrared cameras. *Optical Engineering*, 34(6):1808–1817, 1995.

[21] H. L. Van Trees. *Detection, Estimation and Modulation Theory, Part I.* Wiley, 1968.